

WHAT IS CLAIMED IS:

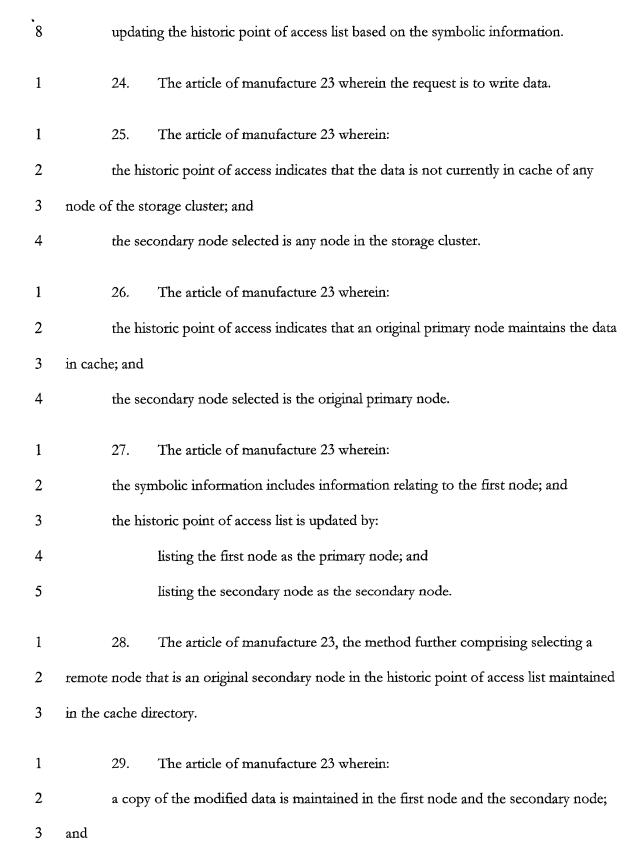
1	1. A method of maintaining cache in a clustered environment comprising:				
2	receiving a request in a primary node of a storage cluster for accessing data;				
3	selecting a secondary node for storing data in cache based on a historic point of				
4	access list maintained in a cache directory;				
5	forwarding modified data and symbolic information to one or more relevant nodes				
6	in the storage cluster; and				
7	updating the historic point of access list based on the symbolic information.				
1	2. The method of claim 1 wherein the request is to write data.				
1	3. The method of claim 1 wherein:				
2	the historic point of access indicates that the data is not currently in cache of any				
3	node of the storage cluster; and				
4	the secondary node selected is any node in the storage cluster.				
1	4. The method of claim 1 wherein:				
2	the historic point of access indicates that an original primary node maintains the data				
3	in cache; and				
4	the secondary node selected is the original primary node.				
1	5. The method of claim 1 wherein:				
2	the symbolic information includes information relating to the first node; and				
3	the historic point of access list is updated by:				
4	listing the first node as the primary node; and				
5	listing the secondary node as the secondary node.				

1	o. The method of claim 1 further comprising selecting a remote hode that is at
2	original secondary node in the historic point of access list maintained in the cache directory
1	7. The method of claim 1 wherein:
2	a copy of the modified data is maintained in the first node and the secondary node;
3	and
4	the symbolic information is maintained in remaining nodes of the storage cluster.
1	8. The method of claim 1 further comprising acquiring a lock on associated
2	tracks on nodes in the storage cluster wherein the locking protocol provides for multiple
3	readers and a single writer.
1	9. The method of claim 1 further comprising:
2	detecting a failure of a node in the storage cluster;
3	broadcasting a failover recovery message to all nodes in the storage cluster; and
4	replicating the data from the primary node or the secondary node to another node i
5	the storage cluster.
1	10. The method of claim 1 further comprising:
2	detecting a failure of a node in the storage cluster;
3	broadcasting a failover recovery message to all nodes in the storage cluster; and
4	destaging the data from the primary node or the secondary node to disk.
1	11. The method of claim 1 further comprising:
2	applying for cluster admission;
3	requesting the symbolic information for new write requests;

4	requesting a modified track list comprising an identifier of modified data and an		
5	associated symbolic entry;		
6	mergin	g the m	nodified track list with any new symbolic entries; and
7	broadc	asting a	availability to remaining nodes in the storage cluster.
1	12.	An ap	paratus for maintaining cache in a clustered environment comprising:
2	(a)	a cach	e;
3	(b)	a cach	e directory comprising a historic point of access list for the cache;
4	(a)	a stora	age node organized in a storage cluster and having an interface for
5	connecting to a host, a storage disk, and one or more additional storage nodes, wherein the		
6	storage node maintains cache and the cache directory, and wherein the storage node is		
7	configured to:		
8		(ï)	receive a request for accessing data;
9		(ii)	select a secondary node for storing data in cache based on the
10	historic	point	of access list;
11		(iii)	forward modified data and symbolic information to one or more
12	additio	nal stoi	rage nodes in the storage cluster; and
13		(iv)	update the historic point of access list based on the symbolic
14	inform	ation.	
1	13.	The ap	pparatus of claim 12 wherein the request is to write data.
1	14.	The ap	oparatus of claim 12 wherein:
2	the his	oric po	pint of access indicates that the data is not currently in cache of the
3	nodes in the storage cluster; and		
4	the sec	ondarv	node selected is any node in the storage cluster.

1	15. The apparatus of claim 12 wherein:
2	the historic point of access indicates that an original primary node maintains the data
3	in cache; and
4	the secondary node selected is the original primary node.
1	16. The apparatus of claim 12 wherein:
2	the symbolic information includes information relating to a first node that receives
3	the request; and
4	the historic point of access list is updated by:
5	listing the first node as the primary node; and
6	listing the secondary node as the secondary node.
1	17. The apparatus of claim 12, wherein the storage node is further configured to
2	select a remote node that is an original secondary node in the historic point of access list
3	maintained in the cache directory.
1	18. The apparatus of claim 12 wherein:
2	a copy of the modified data is maintained in two nodes in the storage cluster; and
3	the symbolic information is maintained in remaining nodes of the storage cluster.
1	19. The apparatus of claim 12, wherein the storage node is further configured to
2	acquire a lock on associated tracks on relevant nodes in the storage cluster wherein the
3	locking protocol provides for multiple readers and a single writer.
1	20. The apparatus of claim 12, wherein the storage node is further configured to
2	detect a failure of a node in the storage cluster;

3	broadcast a failover recovery message to an additional storage node in the storage				
4	cluster; and				
5	replicate the data from one node in the storage cluster to another node in the storage				
6	cluster.				
1	21. The apparatus of claim 12, wherein the storage node is further configured to				
2	detect a failure of a node in the storage cluster;				
3	broadcast a failover recovery message to an additional node in the storage cluster;				
4	and				
5	destage the data from a node in the storage cluster to disk.				
1	22. The apparatus of claim 12 further comprising a new node configured to:				
2	apply for cluster admission;				
3	request the symbolic information for new write requests;				
4	request a modified track list comprising an identifier of modified data and an				
5	associated symbolic entry;				
6	merge the modified track list with any new symbolic entries; and				
7	broadcast availability to remaining nodes in the storage cluster.				
1	23. An article of manufacture, embodying logic to perform a method of				
2	maintaining cache in a clustered environment, the method comprising:				
3	receiving a request in a primary node of a storage cluster for accessing data;				
4	selecting a secondary node for storing data in cache based on a historic point of				
5	access list maintained in a cache directory;				
6	forwarding modified data and symbolic information to one or more relevant nodes				
7	in the storage cluster; and				



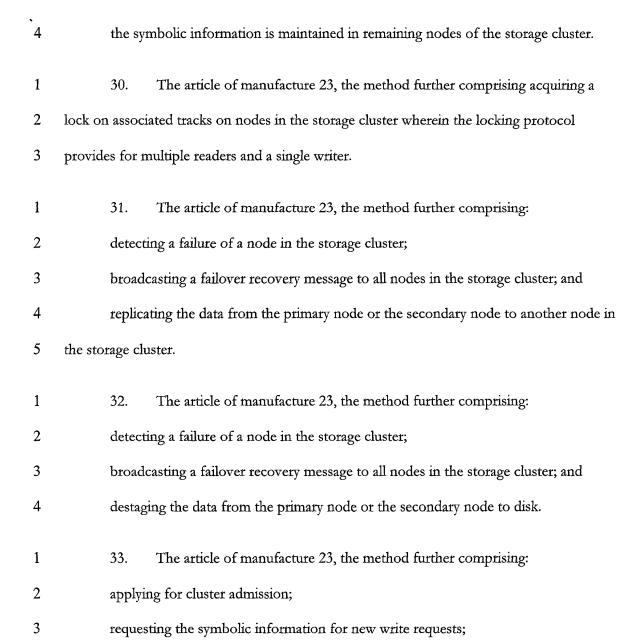
4

5

6

7

associated symbolic entry;



requesting a modified track list comprising an identifier of modified data and an

merging the modified track list with any new symbolic entries; and

broadcasting availability to remaining nodes in the storage cluster.